(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 113935329 A
(43)申请公布日 2022.01.14

(21)申请号 202111192675.7

(22)申请日 2021.10.13

(71)申请人 昆明理工大学
地址 650093 云南省昆明市五华区学府路253号

(72)发明人 郭军军 李岩 余正涛

(74)专利代理机构 昆明人从众知识产权代理有限公司 53204
代理人 何娇

(51)Int.Cl.
G06F 40/30(2020.01)
G06K 9/62(2022.01)
G06N 3/04(2006.01)
G06N 3/08(2006.01)

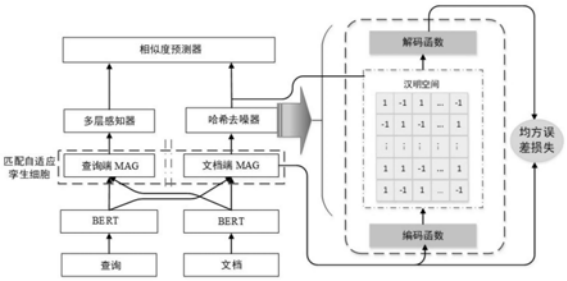
权利要求书3页 说明书12页 附图2页

(54)发明名称

基于自适应特征识别与去噪的非对称文本匹配方法

(57)摘要

本发明涉及基于自适应特征识别与去噪的非对称文本匹配方法，属于自然语言处理技术领域。本发明对于每个非对称文本对，设计成以上下文感知的方式显式识别鉴别性特征并过滤掉不相关的特征。具体地说，首先设计了一个匹配自适应孪生细胞来自适应地识别鉴别特征，从而为每个文本对导出相应的混合表示。然后，提出了一种局部约束哈希去噪器，通过学习一个有区别的低维二进制码来对冗余长文本进行特征级去噪，从而实现更好的相关性学习。在四个不同下游任务的真实数据集上进行的大量实验表明，与最新的最先进的方法相比，本发明方法获得了巨大的性能增益，为后续的信息检索和答案选择等下游任务提供了支撑。



CN 113935329 A

1.基于自适应特征识别与去噪的非对称文本匹配方法,其特征在于:所述方法的具体步骤如下:

Step1、首先将问题-答案匹配数据集和查询-文档匹配数据集进行预处理;

Step2、将Step1预处理过的每个非对称文本对,利用一个基于BERT的上下文编码器对每个非对称文本对进行上下文表示;基于一个自适应匹配孪生细胞来自适应地识别鉴别特征,从而为每个非对称文本对导出相应的混合表示;提出了一种局部约束哈希去噪器,通过学习一个有区别的低维二进制码来对冗余长文本进行特征级去噪;最后利用相似度预测器来获得非对称文本对的匹配分数。

2.根据权利要求1所述的基于自适应特征识别与去噪的非对称文本匹配方法,其特征在于:所述Step1中,问题-答案匹配数据集包括insuranceQA、wikiQA和yahooQA,查询-文档匹配数据集采用MS MARCO;预处理包括将文本中的特殊字符利用正则表达式进行匹配删除。

3.根据权利要求1所述的基于自适应特征识别与去噪的非对称文本匹配方法,其特征在于:所述Step2中,将Step1预处理过的每个非对称文本对,利用一个基于BERT的上下文编码器对每个非对称文本对进行上下文表示包括:

选择使用BERT作为上下文编码器,遵循BERT输入的格式,将特定的标记符号[CLS]放在序列的开头,即,{[CLS],q_1,q_2,…,q_l}和{[CLS],d_1,d_2,…,d_t},这里,基于BERT的上下文编码器描述如下:

U_Q=BERT([CLS],q_1,q_2,…,q_l) (1)

V_D=BERT([CLS],d_1,d_2,…,d_t) (2)

其中,$U_Q \in R^{l\times d}$和$V_D \in R^{t\times d}$分别表示查询Q和文档D的上下文表示;d表示BERT的输出维度;为了减少参数量,防止过拟合,并促进跨文本对的信息交互,查询和文档共享一个上下文编码器。

4.根据权利要求1所述的基于自适应特征识别与去噪的非对称文本匹配方法,其特征在于:所述Step2中,基于一个自适应匹配孪生细胞来自适应地识别鉴别特征,从而为每个非对称文本对导出相应的混合表示包括:

使用自适应匹配孪生细胞称为MAGS来模拟特征识别过程;自适应匹配孪生细胞它是一个具有两个子单元MAG的并行架构,即查询端MAG和文档端MAG;由于查询端和文档端MAG都是相同的,其中,查询端MAG为:

给定提取的查询的上下文表示U_Q=[u_1,…,u_l]及文档的上下文表示V_D=[v_1,…,v_t],l、t分别表示查询文本的长度、文档文本的长度,为了识别鉴别性特征并将其合成为相关性特征;具体而言,首先计算单词级相似度,如下所示:

$$S = U_Q V_D^T \quad (3)$$

其中,$S \in R^{l\times t}$是两个序列中所有单词对的相似性矩阵;然后,将这些相似性分数标准化,并根据V_D为查询Q中的每一个单词推导出参考表示:

R_Q=soft max(S)V_D (4)

这一操作的目的是为了根据S对V_D执行软特征选择;也就是说,文档D中的相关信息被传输到表示Q;

然而，在这个参考表示过程中，Q中的不相关的信息还提供了进一步的相关学习；首先通过考虑参考表示与原始表示的差异来构造补充特征：$D_Q=U_Q-R_Q$；此外，为了识别鉴别性特征，首先利用以S表示的相似模式来识别R_Q、D_Q这两种语义信号中的重要特征，如下所示：

$$E=\sigma(W_1S+B_1) \quad (5)$$

$$F^{(r)}=R_Q\odot E \quad (6)$$

$$F^{(d)}=D_Q\odot(1-E) \quad (7)$$

其中σ(•)表示sigmoid激活函数，W_1和B_1分别为变换矩阵和偏置矩阵，以及⊙是元素按位乘积操作；然后，进一步连接这两个部分即$F_i^{(r)}$和$F_i^{(d)}$，通过类似于公式5的注意机制：

$$p_i=\sigma(W_2S_i+B_2) \quad (8)$$

$$F_i^{(c)}=p_i\bullet F_i^{(r)}\oplus(1-p_i)\bullet F_i^{(d)} \quad (9)$$

其中，S_i，$F_i^{(r)}$和$F_i^{(d)}$分别对应矩阵S、$F^{(r)}$、$F^{(d)}$的第i行，符号⊕是向量串联操作，d表示BERT的输出维度，W_2和B_2也分别为变换矩阵和偏置矩阵；然后，采用一个高速网络来生成每个单词的鉴别特征h_i^Q；

$$p_i=\mathrm{relu}(W_3F_i^{(c)}+b_3) \quad (10)$$

$$g_i=\mathrm{sigmoid}(W_4F_i^{(c)}+b_4) \quad (11)$$

$$i_i=(1-g_i)\odot F_i^{(c)}+g_i\odot p_i \quad (12)$$

$$h_i^Q=W_5i_i+b_5 \quad (13)$$

其中，W_3，$W_4\in R^{2d\times 2d}$和$W_5\in R^{d\times 2d}$表示参数矩阵，b_3，b_4，b_5表示偏置向量；将合成的混合鉴别特征形成矩阵：$H^Q=[h_1^Q;\cdots;h_l^Q]$。

5.根据权利要求1所述的基于自适应特征识别与去噪的非对称文本匹配方法，其特征在于：所述Step2中，提出了一种局部约束哈希去噪器，通过学习一个有区别的低维二进制码来对冗余长文本进行特征级去噪具体包括：

局部约束哈希去噪器定义了编码函数F_{en}，一个哈希函数F_h，以及解码函数F_{de}；

(1)编码函数F_{en}映射表示形式H^D转化为低维矩阵$B\in R^{t\times h}$；这里，使用一个由三层多层感知机MLP实现的前馈神经网络FNN(•)来为F_{en}建模；此外，为了过滤语义噪声和缓解梯度消失问题，选择使用relu(•)作为第二层的激活功能，它能跳过不必要的特征并保留鉴别线索；编码过程总结如下：

$$B=F_{en}(H^D)=\mathrm{FFN}(H^D) \quad (14)$$

(2)哈希函数F_h被用来学习有区别的二元矩阵表示，以达到净化和有效匹配的目的；sgn(•)函数是二值化的最佳选择，但是sgn(•)是不可微的；因此，使用一个近似函数tanh(•)替换sgn(•)用于支持模型训练；具体而言，哈希函数表示如下；

$$B^D=F_h(B)=\tanh(\alpha B) \quad (15)$$

请注意，引入超参数α是为了使哈希函数更加灵活，并生成平衡的、有区别的哈希码，为了确保B中的值属于{-1,1}，定义了一个额外的约束：

$$L_1=\left\|B^D-B^{(b)}\right\|_F^2 \quad (16)$$

其中$B^{(b)}=\mathrm{sgn}(B)$表示H^D的二进制矩阵表示，$||\ ||_F$表示F-范数，B^D为文档D经过哈希去噪器之后的上下文表示，也就是哈希函数生成的二值码；

(3)解码函数F_{de}从B^D中重构了H^D.它由三层多层感知器组成,用于解码二进制矩阵B^D回到原来的那个H^D,因此,重构序列矩阵H_r^D定义如下:

$$H_r^D = F_{de}(B^D) = FFN^T(B^D) \quad (17)$$

其中$FNN^T(\cdot)$是解码器函数,为了减少重构过程中语义的丢失,增加了均方误差MSE(·)作为训练模型时的约束条件;

$$L_2 = MSE(H_r^D, H^D) = \frac{1}{t \times d}\left\| H_r^D - H^D \right\|_F^2 \quad (18)$$

需要强调的是,还为H^Q执行哈希去噪,使用单个MLP层更新查询Q的矩阵表示H^Q,以匹配哈希去噪器的维数,即h;

$H^Q = MLP(H^Q)$ (19)。

6.根据权利要求1所述的基于自适应特征识别与去噪的非对称文本匹配方法,其特征在于:所述Step2中,利用相似度预测器来获得非对称文本对的匹配分数包括:

对于查询Q经过哈希去噪器之后的上下文表示$H^Q = [h_1^Q; \cdots; h_l^Q]$和文档D经过哈希去噪器之后的上下文表示$B^D = [b_1^D; \cdots; b_t^D]$,查询Q和文档D之间的匹配分数G(Q,D)通过MaxSim运算符进行估计,如下所示:

$$G(Q,D) = \sum_i^l \max_j^t Norm(h_i^Q) \bullet (Norm(b_j^D))^T \quad (20)$$

其中Norm(·)表示L2规范化,这样,当计算任意两个隐藏表示的内积时,结果在[-1,1],即,相当于其余弦相似性,h_i^Q是H^Q中的第i个词的向量表示,b_j^D是B^D的第j个向量表示。

7.根据权利要求5所述的基于自适应特征识别与去噪的非对称文本匹配方法,其特征在于:所述Step2中,模型优化包括:

在训练阶段,通过基于三重铰链损失使用负采样策略:

$L_3 = \max\{0, 0.1 - G(Q,D) + G(Q,D^-)\}$ (21)

其中D^-是从训练集中取样的相应负样本文档,G(Q,D)是查询Q和文档D之间的匹配分数;

最后,结合铰链损失和哈希去噪器中两个约束;也就是说,最终优化目标是L_1、L_2和L_3的线性融合:

$$\begin{aligned} \min_\theta L &= \sum_{(Q,D,D^-)} [L_3 + \delta \bullet L_1 + \gamma \bullet L_2] \\ &= \sum_{(Q,D,D^-)} [\max(0, 0.1 - G(Q,D) + G(Q,D^-)) + \delta \bullet L_1 + \gamma \bullet L_2 \\ &\quad \delta \bullet (\left\| B^D - B^D \right\|_F^2 + \left\| B^{D^-} - B^{D^-} \right\|_F^2) + \\ &\quad \gamma \bullet (MSE(H_r^D, H^D) + MSE(H_r^{D^-}, H^{D^-}))] \end{aligned} \quad (22)$$

其中δ和γ是可调超参数,它们分别控制两个约束的重要性,θ是参数集,使用Adam在小批量上以端到端的方式更新参数,B^D是文档D经过哈希去噪器之后的上下文表示,也就是哈希函数生成的二值码,B^D表示文档D利用sgn符号函数生成的哈希码。

基于自适应特征识别与去噪的非对称文本匹配方法

技术领域

[0001] 本发明涉及基于自适应特征识别与去噪的非对称文本匹配方法,属于自然语言处理技术领域。

背景技术

[0002] 文本匹配(TM)是信息检索和自然语言处理领域中一项有价值但具有挑战性的任务。给定一对文档,TM旨在预测它们的语义关系。请注意,在许多信息检索系统,问答系统和对话系统中,高效的匹配算法是不可或缺的资产。在大多数应用场景中,匹配的序列对(例如,查询文档、关键字文档和问答对)通常在信息量上存在很大差异(例如,不对称文本匹配)。例如,匹配对中的两个文档在InsuranceQA数据集中的平均字数为7.15和95.54(即数量级)。短查询和长文档的不对称性使得它成为一项非常重要的任务。非对称文本匹配已成为许多下游任务(如信息检索和自然语言处理)日益增长的需求。这里,不对称意味着匹配所涉及的文档包含不同数量的信息,例如,针对相对较长的文档的简短查询。

[0003] 早期的解决方案可分为两类,即基于表示的模型和基于交互的模型。前一种解决方案利用递归神经网络(RNN)和长短期记忆网络(LSTM)通过独立处理每个文档来学习文档对的潜在表示,包括DSSM、SNRM和ARC-I。相比之下,后者捕获了它们之间细粒度的交互信号。人们普遍认为,利用交互信号可以极大地提高关联学习能力。示例包括DRMM、KNR和ARC-II。最近,随着像BERT这样的深度预训练语言模型(LMs)的出现,最新的基于LMs的深度关联模型极大地推动了最新技术的发展。具体来说,LMs在大规模语料库上进行预训练,然后通过计算句子对的上下文语义表示将其应用于TM任务。目标是进一步消除文档和查询之间的词汇不匹配。尽管这些努力取得了显著的性能提升,但这些模型的主要缺点是在不对称文本之间忽略了进一步的特征识别和去噪,这可能有助于提高匹配性能。

发明内容

[0004] 本发明提供了基于自适应特征识别与去噪的非对称文本匹配方法,设计了一个匹配自适应孪生细胞(MAGS)来自适应地识别鉴别特征,从而为每个文本对导出相应的混合表示,还提出了一种局部约束哈希去噪器,通过学习一个有区别的低维二进制码来对冗余长文本进行特征级去噪,从而实现更好的相关性学习。

[0005] 本发明的技术方案是:基于自适应特征识别与去噪的非对称文本匹配方法,所述方法的具体步骤如下:

[0006] Step1、首先将问题-答案匹配数据集和查询-文档匹配数据集进行预处理;

[0007] Step2、将Step1预处理过的每个非对称文本对,利用一个基于BERT的上下文编码器对每个非对称文本对进行上下文表示;基于一个自适应匹配孪生细胞来自适应地识别鉴别特征,从而为每个非对称文本对导出相应的混合表示;提出了一种局部约束哈希去噪器,通过学习一个有区别的低维二进制码来对冗余长文本进行特征级去噪;最后利用相似度预测器来获得非对称文本对的匹配分数。

[0008] 作为本发明的进一步方案，所述Step1中，问题-答案匹配数据集包括insuranceQA、wikiQA和yahooQA，查询-文档匹配数据集采用MS MARCO；预处理包括将文本中的特殊字符利用正则表达式进行匹配删除。

[0009] 作为本发明的进一步方案，所述Step2中，将Step1预处理过的每个非对称文本对，利用一个基于BERT的上下文编码器对每个非对称文本对进行上下文表示包括：

[0010] 选择使用BERT作为上下文编码器，遵循BERT输入的格式，将特定的标记符号[CLS]放在序列的开头，即，{[CLS]，q_1，q_2，…，q_l}和{[CLS]，d_1，d_2，…，d_t}，这里，基于BERT的上下文编码器描述如下：

[0011] U_Q＝BERT([CLS]，q_1，q_2，…，q_l) (1)

[0012] V_D＝BERT([CLS]，d_1，d_2，…，d_t) (2)

[0013] 其中，$U_Q \in R^{l\times d}$和$V_D \in R^{t\times d}$分别表示查询Q和文档D的上下文表示；d表示BERT的输出维度；为了减少参数量，防止过拟合，并促进跨文本对的信息交互，查询和文档共享一个上下文编码器。

[0014] 作为本发明的进一步方案，所述Step2中，基于一个自适应匹配孪生细胞来自适应地识别鉴别特征，从而为每个非对称文本对导出相应的混合表示包括：

[0015] 使用自适应匹配孪生细胞称为MAGS来模拟特征识别过程；自适应匹配孪生细胞它是一个具有两个子单元MAG的并行架构，即查询端MAG和文档端MAG；由于查询端和文档端MAG都是相同的，其中，查询端MAG为：

[0016] 给定提取的查询的上下文表示U_Q＝[u_1，…，u_l]及文档的上下文表示V_D＝[v_1，…，v_t]，l、t分别表示查询文本的长度、文档文本的长度，为了识别鉴别性特征并将其合成为相关性特征；具体而言，首先计算单词级相似度，如下所示：

[0017] $$S = U_Q V_D^T \quad (3)$$

[0018] 其中，$S \in R^{l\times t}$是两个序列中所有单词对的相似性矩阵；然后，将这些相似性分数标准化，并根据V_D为查询Q中的每一个单词推导出参考表示：

[0019] R_Q＝softmax(S)V_D (4)

[0020] 这一操作的目的是为了根据S对V_D执行软特征选择；也就是说，文档D中的相关信息被传输到表示Q；

[0021] 然而，在这个参考表示过程中，Q中的不相关的信息还提供了进一步的相关学习；首先通过考虑参考表示与原始表示的差异来构造补充特征：D_Q＝U_Q-R_Q；此外，为了识别鉴别性特征，首先利用以S表示的相似模式来识别R_Q、D_Q这两种语义信号中的重要特征，如下所示：

[0022] E＝σ(W_1S+B_1) (5)

[0023] $F^{(r)}$＝R_Q⊙E (6)

[0024] $F^{(d)}$＝D_Q⊙(1-E) (7)

[0025] 其中σ(•)表示sigmoid激活函数，W_1和B_1分别为变换矩阵和偏置矩阵，以及⊙是元素按位乘积操作；然后，进一步连接这两个部分即$F_i^{(r)}$和$F_i^{(d)}$，通过类似于公式5的注意机制：

[0026] p_i＝σ(W_2S_i+B_2) (8)

[0027] $F_i^{(c)} = p_i \bullet F_i^{(r)} \oplus (1-p_i) \bullet F_i^{(d)}$ (9)

[0028] 其中，S_i，$F_i^{(r)}$和$F_i^{(d)}$分别对应矩阵S、$F^{(r)}$、$F^{(d)}$的第i行，符号⊕是向量串联操作，d表示BERT的输出维度，W_2和B_2也分别为变换矩阵和偏置矩阵；然后，采用一个高速网络来生成每个单词的鉴别特征h_i^Q；

[0029] $p_i = relu(W_3 F_i^{(c)} + b_3)$ (10)

[0030] $g_i = sig\ moid(W_4 F_i^{(c)} + b_4)$ (11)

[0031] $i_i = (1-g_i) \odot F_i^{(c)} + g_i \odot p_i$ (12)

[0032] $h_i^Q = W_5 i_i + b_5$ (13)

[0033] 其中，W_3，$W_4 \in R^{2d \times 2d}$和$W_5 \in R^{d \times 2d}$表示参数矩阵，b_3，b_4，b_5表示偏置向量；将合成的混合鉴别特征形成矩阵：$H^Q = [h_1^Q; \cdots; h_l^Q]$。

[0034] 作为本发明的进一步方案，所述Step2中，提出了一种局部约束哈希去噪器，通过学习一个有区别的低维二进制码来对冗余长文本进行特征级去噪具体包括：

[0035] 局部约束哈希去噪器定义了编码函数F_{en}，一个哈希函数F_h，以及解码函数F_{de}；(1)编码函数F_{en}映射表示形式H^D转化为低维矩阵$B \in R^{t \times h}$；这里，使用一个由三层多层感知机MLP实现的前馈神经网络FNN(•)来为F_{en}建模；此外，为了过滤语义噪声和缓解梯度消失问题，选择使用relu(•)作为第二层的激活功能，它能跳过不必要的特征并保留鉴别线索；编码过程总结如下：

[0036] $B = F_{en}(H^D) = FFN(H^D)$ (14)

[0037] (2)哈希函数F_h被用来学习有区别的二元矩阵表示，以达到净化和有效匹配的目的；sgn(•)函数是二值化的最佳选择，但是sgn(•)是不可微的；因此，使用一个近似函数tanh(•)替换sgn(•)用于支持模型训练；具体而言，哈希函数表示如下；

[0038] $B^D = F_h(B) = \tanh(\alpha B)$ (15)

[0039] 请注意，引入超参数α是为了使哈希函数更加灵活，并生成平衡的、有区别的哈希码，为了确保B中的值属于{-1,1}，定义了一个额外的约束：

[0040] $L_1 = \left\| B^D - B^{(b)} \right\|_F^2$ (16)

[0041] 其中$B^{(b)} = sgn(B)$表示H^D的二进制矩阵表示，|| ||$_F$表示F-范数，B^D为文档D经过哈希去噪器之后的上下文表示，也就是哈希函数生成的二值码；

[0042] (3)解码函数F_{de}从B^D中重构了H^D.它由三层多层感知器组成，用于解码二进制矩阵B^D回到原来的那个H^D，因此，重构序列矩阵H_r^D定义如下：

[0043] $H_r^D = F_{de}(B^D) = FFN^T(B^D)$ (17)

[0044] 其中FNN^T(•)是解码器函数，为了减少重构过程中语义的丢失，增加了均方误差MSE(•)作为训练模型时的约束条件；

[0045] $L_2 = MSE(H_r^D, H^D) = \frac{1}{t \times d} \left\| H_r^D - H^D \right\|_F^2$ (18)

[0046] 需要强调的是，还为H^Q执行哈希去噪，使用单个MLP层更新查询Q的矩阵表示H^Q，以匹配哈希去噪器的维数，即h；

[0047] $H^Q=MLP(H^Q)$ (19)。

[0048] 作为本发明的进一步方案，所述Step2中，利用相似度预测器来获得非对称文本对的匹配分数包括：

[0049] 对于查询Q经过哈希去噪器之后的上下文表示$H^Q=[h_1^Q;\cdots;h_l^Q]$和文档D经过哈希去噪器之后的上下文表示$B^D=[b_1^D;\cdots;b_t^D]$，查询Q和文档D之间的匹配分数G(Q,D)通过MaxSim运算符进行估计，如下所示：

[0050] $$G(Q,D)=\sum_{i}^{l}\max_{j}^{t} Norm(h_i^Q)\bullet(Norm(b_j^D))^T \quad (20)$$

[0051] 其中Norm(•)表示L2规范化，这样，当计算任意两个隐藏表示的内积时，结果在[-1,1]，即，相当于其余弦相似性，h_i^Q是H^Q中的第i个词的向量表示，b_j^D是B^D的第j个向量表示。

[0052] 作为本发明的进一步方案，所述Step2中，模型优化包括：

[0053] 在训练阶段，通过基于三重铰链损失使用负采样策略：

[0054] $L_3=\max\{0,0.1-G(Q,D)+G(Q,D^-)\}$ (21)

[0055] 其中D^-是从训练集中取样的相应负样本文档，G(Q,D)是查询Q和文档D之间的匹配分数；

[0056] 最后，结合铰链损失和哈希去噪器中两个约束；也就是说，最终优化目标是L_1、L_2和L_3的线性融合：

[0057] $$\begin{aligned}\min_{\theta}\ L&=\sum_{(Q,D,D^-)}[L_3+\delta\bullet L_1+\gamma\bullet L_2]\\&=\sum_{(Q,D,D^-)}[\max(0,0.1-G(Q,D)+G(Q,D^-))+\delta\bullet L_1+\gamma\bullet L_2\\&\quad\delta\bullet(\left\|B^D-B^D\right\|_F^2+\left\|B^{D^-}-B^{D^-}\right\|_F^2)+\\&\quad\gamma\bullet(MSE(H_r^D,H^D)+MSE(H_r^{D^-},H^{D^-}))]\end{aligned} \quad (22)$$

[0058] 其中δ和γ是可调超参数，它们分别控制两个约束的重要性，θ是参数集，使用Adam在小批量上以端到端的方式更新参数，B^D是文档D经过哈希去噪器之后的上下文表示，也就是哈希函数生成的二值码，B^D表示文档D利用sgn符号函数生成的哈希码。

[0059] 本发明的有益效果是：

[0060] 本发明对于每个非对称文本对，以上下文感知的方式显式区分区别性特征并过滤掉不相关的特征；具体地说，首先设计了一个匹配自适应孪生细胞(MAGS)来自适应地识别鉴别性特征，从而为每个文本对导出相应的混合表示。然后，本发明进一步提出了一种局部约束哈希去噪器，通过学习一个有区别的低维二进制码来对冗余长文本进行特征级去噪，从而实现更好的相关性学习。在四个不同下游任务的真实数据集上进行的大量实验表明，与最新的最先进的替代方案相比，所提出的本发明获得了巨大的性能增益。

附图说明

[0061] 图1为本发明中的模型示意图；

[0062] 图2为本发明自适应匹配孪生细胞结构图；

[0063] 图3为本发明超参数灵敏性分析的折线图。

具体实施方式

[0064] 实施例1:如图1-3所示,基于自适应特征识别与去噪的非对称文本匹配方法,所述方法的具体步骤如下:

[0065] Step1、首先将问题-答案匹配数据集和查询-文档匹配数据集进行预处理;

[0066] Step1.1、将问题-答案匹配数据集(insuranceQA、wikiQA和yahooQA)和查询-文档匹配数据集(MS MARCO)进行预处理;将文本中的特殊字符利用正则表达式进行匹配删除。其中,查询-文档匹配数据集(MS MARCO)是一个包含880万个网页段落的集合,包含大约4亿个查询,正面和负面段落组成的元组。本发明报告了MSMARCO Dev集合的结果,该集合包含大约6900个查询;问题-答案匹配数据集规模如表1所示:

[0067] 表1为QA数据集的统计信息(insuranceQA测试集包括Test1和Test2)

[0068]

数据集	insuranceQA	wikiQA	yahooQA
问题数量(Train)	12887	873	50112
问题数量(Dev)	1000	126	6289
问题数量(Test)	1800/1800	243	6283
每个问题的候选答案数量	500	9	5

[0069] Step2、将Step1预处理过的每个非对称文本对,利用一个基于BERT的上下文编码器对每个非对称文本对进行上下文表示;基于一个自适应匹配孪生细胞来自适应地识别鉴别特征,从而为每个非对称文本对导出相应的混合表示;提出了一种局部约束哈希去噪器,通过学习一个有区别的低维二进制码来对冗余长文本进行特征级去噪;最后利用相似度预测器来获得非对称文本对的匹配分数。

[0070] 作为本发明的进一步方案,所述Step2中,将Step1预处理过的每个非对称文本对,利用一个基于BERT的上下文编码器对每个非对称文本对进行上下文表示包括:

[0071] 选择使用BERT作为上下文编码器,遵循BERT输入的格式,将特定的标记符号[CLS]放在序列的开头,即,{[CLS],q_1,q_2,…,q_l}和{[CLS],d_1,d_2,…,d_t},这里,基于BERT的上下文编码器描述如下:

[0072] U_Q=BERT([CLS],q_1,q_2,…,q_l) (1)

[0073] V_D=BERT([CLS],d_1,d_2,…,d_t) (2)

[0074] 其中,$U_Q \in R^{l\times d}$和$V_D \in R^{t\times d}$分别表示查询Q和文档D的上下文表示;d表示BERT的输出维度;为了减少参数量,防止过拟合,并促进跨文本对的信息交互,查询和文档共享一个上下文编码器。

[0075] 作为本发明的进一步方案,所述Step2中,基于一个自适应匹配孪生细胞来自适应地识别鉴别特征,从而为每个非对称文本对导出相应的混合表示包括:

[0076] 人类可以一目了然地识别两个序列(例如,查询-文档、关键字-文档和问题-答案)之间的关系。例如,一个训练有素的研究生可以很容易地根据标题和摘要对他/她的研究方向的论文进行分类,因为他/她可以潜意识地识别有区别的特征,而忽略不相关的特征进行决策推理。

[0077] 使用自适应匹配孪生细胞(称为MAGS)来模拟特征识别过程;自适应匹配孪生细胞它是一个具有两个子单元MAG的并行架构,即查询端MAG和文档端MAG;由于查询端和文档端MAG都是相同的,为了简单起见,本发明主要描述查询端MAG(即图2说明了整体架构),其中,

查询端MAG为：

[0078] 给定提取的查询的上下文表示$U_Q=[u_1,\cdots,u_l]$及文档的上下文表示$V_D=[v_1,\cdots,v_t]$，l、t分别表示查询文本的长度、文档文本的长度，为了识别鉴别性特征并将其合成为相关性特征；具体而言，首先计算单词级相似度，如下所示：

[0079] $$S=U_Q V_D^T \tag{3}$$

[0080] 其中，$S\in R^{l\times t}$是两个序列中所有单词对的相似性矩阵；然后，将这些相似性分数标准化，并根据V_D为查询Q中的每一个单词推导出参考表示：

[0081] $$R_Q=\mathrm{softmax}(S)V_D \tag{4}$$

[0082] 这一操作的目的是为了根据S对V_D执行软特征选择；也就是说，文档D中的相关信息被传输到表示Q；

[0083] 然而，在这个参考表示过程中，Q中的不相关的信息还提供了进一步的相关学习；首先通过考虑参考表示与原始表示的差异来构造补充特征：$D_Q=U_Q-R_Q$；此外，为了识别鉴别性特征，首先利用以S表示的相似模式来识别R_Q、D_Q这两种语义信号中的重要特征，如下所示：

[0084] $$E=\sigma(W_1S+B_1) \tag{5}$$

[0085] $$F^{(r)}=R_Q\odot E \tag{6}$$

[0086] $$F^{(d)}=D_Q\odot(1-E) \tag{7}$$

[0087] 其中σ(•)表示sigmoid激活函数，W_1和B_1分别为变换矩阵和偏置矩阵，以及⊙是元素按位乘积操作；然后，进一步连接这两个部分即$F_i^{(r)}$和$F_i^{(d)}$，通过类似于公式5的注意机制：

[0088] $$p_i=\sigma(W_2S_i+B_2) \tag{8}$$

[0089] $$F_i^{(c)}=p_i\bullet F_i^{(r)}\oplus(1-p_i)\bullet F_i^{(d)} \tag{9}$$

[0090] 其中，S_i，$F_i^{(r)}$和$F_i^{(d)}$分别对应矩阵S、$F^{(r)}$、$F^{(d)}$的第i行，符号⊕是向量串联操作，d表示BERT的输出维度，W_2和B_2也分别为变换矩阵和偏置矩阵；然后，采用一个高速网络来生成每个单词的鉴别特征h_i^Q；

[0091] $$p_i=\mathrm{relu}(W_3F_i^{(c)}+b_3) \tag{10}$$

[0092] $$g_i=\mathrm{sig\ moid}(W_4F_i^{(c)}+b_4) \tag{11}$$

[0093] $$i_i=(1-g_i)\odot F_i^{(c)}+g_i\odot p_i \tag{12}$$

[0094] $$h_i^Q=W_5i_i+b_5 \tag{13}$$

[0095] 其中，W_3，$W_4\in R^{2d\times 2d}$和$W_5\in R^{d\times 2d}$表示参数矩阵，b_3，b_4，b_5表示偏置向量；将合成的混合鉴别特征形成矩阵：$H^Q=[h_1^Q;\cdots;h_l^Q]$。

[0096] 文档端MAG：与查询端MAG类似，文档侧MAG单元为同一流程切换Q和D的角色，但是两个子单元的参数是不共享的。本发明使用$H^D=[h_1^D;\cdots;h_t^D]$表示由文档侧MAG导出的鉴别特征。

[0097] 作为本发明的进一步方案，所述Step2中，提出了一种局部约束哈希去噪器，通过学习一个有区别的低维二进制码来对冗余长文本进行特征级去噪具体包括：

[0098] 由于文档D比查询Q大得多,所以由文档端MAG执行的鉴别特征提取仍然会引入许多语义噪声。在这里,本发明采用了一种局部约束哈希去噪器来进一步过滤掉不相关的特征。更具体地说,局部约束哈希去噪器定义了编码函数F_{en},一个哈希函数F_h,以及解码函数F_{de};

[0099] (1)编码函数F_{en}映射表示形式H^D转化为低维矩阵$B\in R^{t\times h}$;这里,使用一个由三层多层感知机MLP实现的前馈神经网络FNN(•)来为F_{en}建模;此外,为了过滤语义噪声和缓解梯度消失问题,选择使用relu(•)作为第二层的激活功能(其他为tanh(•)),它能跳过不必要的特征并保留鉴别线索;编码过程总结如下:

[0100] $B=F_{en}(H^D)=FFN(H^D)$ (14)

[0101] (2)哈希函数F_h被用来学习有区别的二元矩阵表示,以达到净化和有效匹配的目的;sgn(•)函数是二值化的最佳选择,但是sgn(•)是不可微的;因此,使用一个近似函数tanh(•)替换sgn(•)用于支持模型训练;具体而言,哈希函数表示如下;

[0102] $B^D=F_h(B)=\tanh(\alpha B)$ (15)

[0103] 请注意,引入超参数α是为了使哈希函数更加灵活,并生成平衡的、有区别的哈希码,为了确保B中的值属于{-1,1},定义了一个额外的约束:

[0104] $$L_1=\left\|B^D-B^{(b)}\right\|_F^2 \quad (16)$$

[0105] 其中$B^{(b)}=sgn(B)$表示H^D的二进制矩阵表示,$||\ ||_F$表示F-范数,B^D为文档D经过哈希去噪器之后的上下文表示,也就是哈希函数生成的二值码;

[0106] (3)解码函数F_{de}从B^D中重构了H^D.它由三层多层感知器组成,用于解码二进制矩阵B^D回到原来的那个H^D,因此,重构序列矩阵H_r^D定义如下:

[0107] $$H_r^D=F_{de}(B^D)=FFN^T(B^D) \quad (17)$$

[0108] 其中$FNN^T(•)$是解码器函数,为了减少重构过程中语义的丢失,增加了均方误差MSE(•)作为训练模型时的约束条件;

[0109] $$L_2=MSE(H_r^D,H^D)=\frac{1}{t\times d}\left\|H_r^D-H^D\right\|_F^2 \quad (18)$$

[0110] 需要强调的是,还为H^Q执行哈希去噪,使用单个MLP层更新查询Q的矩阵表示H^Q,以匹配哈希去噪器的维数,即h;

[0111] $H^Q=MLP(H^Q)$ (19)。

[0112] 作为本发明的进一步方案,所述Step2中,利用相似度预测器来获得非对称文本对的匹配分数包括:

[0113] 对于查询Q经过哈希去噪器之后的上下文表示$H^Q=[h_1^Q;\cdots;h_l^Q]$和文档D经过哈希去噪器之后的上下文表示$B^D=[b_1^D;\cdots;b_t^D]$,查询Q和文档D之间的匹配分数G(Q,D)通过MaxSim运算符进行估计,如下所示:

[0114] $$G(Q,D)=\sum_i^l \max_j^t Norm(h_i^Q)\bullet(Norm(b_j^D))^T \quad (20)$$

[0115] 其中Norm(•)表示L2规范化,这样,当计算任意两个隐藏表示的内积时,结果在[-

1,1]，即，相当于其余弦相似性，h_i^Q是H^Q中的第i个词的向量表示，b_j^D是B^D的第j个向量表示。

[0116] 作为本发明的进一步方案，所述Step2中，模型优化的目的是指导ADDAX的相关学习，帮助估计不对称文本对的匹配分数，模型优化包括：

[0117] 在训练阶段，通过基于三重铰链损失使用负采样策略：

[0118] $L_3=\max\{0,0.1-G(Q,D)+G(Q,D^-)\}$ (21)

[0119] 其中D^-是从训练集中取样的相应负样本文档，G(Q,D)是查询Q和文档D之间的匹配分数；

[0120] 最后，结合铰链损失和哈希去噪器中两个约束；也就是说，最终优化目标是L_1、L_2和L_3的线性融合：

[0121]
$$\begin{aligned}\min_{\theta} L &= \sum_{(Q,D,D^-)} [L_3+\delta\bullet L_1+\gamma\bullet L_2] \\ &= \sum_{(Q,D,D^-)} [\max(0,0.1-G(Q,D)+G(Q,D^-))+\delta\bullet L_1+\gamma\bullet L_2 \\ &\quad \delta\bullet(\left\|B^D-B^D\right\|_F^2+\left\|B^{D^-}-B^{D^-}\right\|_F^2)+ \\ &\quad \gamma\bullet(MSE(H_r^D,H^D)+MSE(H_r^{D^-},H^{D^-}))]\end{aligned} \tag{22}$$

[0122] 其中δ和γ是可调超参数，它们分别控制两个约束的重要性，θ是参数集，使用Adam在小批量上以端到端的方式更新参数，B^D是文档D经过哈希去噪器之后的上下文表示，也就是哈希函数生成的二值码，B^D表示文档D利用sgn符号函数生成的哈希码。

[0123] 为了验证本发明的有效性，以下介绍评价指标、实验的详细参数设置及对比的基准模型，并对实验结果进行分析和讨论，本发明提出了一种新的用于非对称文本匹配的自适应特征识别和去噪模型，称为ADDAX。

[0124] 1.本发明的评价指标主要采用MRR(Mean Reciprocal Rank)、P@1(Precision at 1)、MAP(Mean Average Precision)。在本发明的实验中，本发明选择$BERT_{base}$作为ADDAX中的上下文编码器。更具体地说，本发明设置了隐藏维度h＝300。insuranceQA、wikiQA、yahooQA和MS MARCO的最小批量大小分别设置为32、32、64和64。随机失活率设置为0.1。insuranceQA、MS MARCO、wikiQA和yahooQA的学习率分别为5e-6，5e-6，1e-5和9e-6。insuranceQA的训练次数为60次，wikiQA的训练次数为18次，yahooQA的训练次数为9次。此外，本发明在MS MARCO迭代次数为20万。α、δ和γ的值分别设置为5、1e-6、0.003。

[0125] 2.由于非对称文本匹配已成为许多下游任务(如信息检索和答案选择)日益增长的需求，所以在四个真实数据集上进行实验，以评估本发明提出的ADDAX的有效性，包括问答和文档检索任务。同时，本发明将ADDAX与两种最先进的基线进行比较。第一种类型可以执行问答匹配，另一种类型可以执行文档检索。

[0126] 问答匹配：选择用于答案选择的基线模型可分为四类：(a)传统单一模型：IARNN-GATE，AP-CNN，RNN-POA，AP-BiLSTM，HD-LSTM，AP-LSTM，Multihop-Sequential-LSTM，HyperQA，MULT，TFM+HN，LSTM-CNN+HN；(b)融入外部知识的单一模型：KAN，CKANN；(c)集合模型：$SUM_{BASE,PTK}$，LRXNET，SD(BiLSTM+TFM)；(d)基于BERT的模型：HAS，BERT-pooling和BERT-attention。

[0127] 文档检索：本发明首先将BM25作为基线，这是有代表性的常规检索方法。包括基于

交互的神经排序模型，如KNRM、fastText+ConvKNRM和Duet。此外，由于提出的ADDAX采用BERT作为上下文编码器，因此本发明选择了几种最新的基于预训练语言模型的方法，包括$BERT_{base}$ranker、DeepCT、docT5query、ColBERT、TCTColBERT、COIL-tok和COIL-full。此外，本发明还增加了两个密集检索器用于性能比较，即CLEAR和ADORE+STAR。

[0128] 3.为了验证本发明提出的ADDAX的有效性，并考虑到不同的任务性质和数据特征，四个数据集中现有的最先进模型是完全不同的。表2总结了在问答匹配相应的三个数据集上选择答案的22种方法的性能。本发明选择在每个数据集上分别讨论实验结果。

[0129] 表2为在QA数据集上，本发明提出的ADDAX和几个最先进的基线之间的性能比较，不适用的结果用“--”表示，没有可用的“不适用”。最佳结果以粗体突出显示。

[0130]

	insuranceQA		wikiQA		yahooQA	
模型	P@1(Test1)	P@1(Test2)	MAP	MRR	P@1	MRR
IARNN-GATE	70.10	62.80	72.58	73.94	72.60	73.90
AP-CNN	69.80	66.30	68.86	69.57	56.00	72.60
AP-BiLSTM	71.70	66.40	67.05	68.42	56.80	73.10
HD-LSTM	--	--	--	--	55.70	73.50
HyperQA	n.a.	n.a	71.20	72.70	68.30	80.10
RNN-POA	n.a	n.a	72.12	73.12	n.a	n.a.
Multihop-Sequential-LSTM	70.50	66.90	72.20	73.80	n.a	n.a
AP-LSTM	n.a	n.a	69.00	64.80	68.90	69.60
MULT	75.20	73.40	74.33	75.45	n.a	n.a
LSTM-CNN+HN	73.30	69.10	--	--	--	--
TFM+HN	75.60	73.40	--	--	--	--
KAN(Tgt-Only)	71.50	68.80	--	--	67.20	80.30
KAN	75.20	72.50			74.40	84.00
CKANN	76.30	75.10	73.20	75.50	84.40	90.20
CKANN-L	75.90	**74.90**	72.80	73.90	84.20	90.60
$SUM_{BASE,PTK}$	--	--	75.59	77.00	--	--
LRXNET	--	--	76.57	75.10	--	--
SD(BiLSTM)	--	--	70.40	71.20	--	--
BERT-pooling	74.52	71.97	77.22	78.27	73.49	81.93
BERT-attention	76.12	74.12	80.65	81.63	74.78	82.68
HAS	75.94	73.39	81.01	82.22	74.15	82.28
ADDAX	**77.83**	74.83	**82.50**	**83.38**	**87.63**	**90.69**

[0131]

[0132] insuranceQA的结果。表2总结了insuranceQA数据集的实验结果。本发明观察到传统的单一模型，如AP-CNN、AP-BiLSTM、Multihop-Sequential LSTM和IARNN-GATE在两个测试集的P@1值比MULT、LSTM-CNN+HN和TFM+HN低得多。此外，与单一模型相比，基于BERT的方法(例如，BERT-pooling、BERT-attention和HAS)始终产生更好的性能，这并不奇怪。因为BERT是在大规模语言语料库上预先训练的，它可以利用丰富的公共知识来帮助消除词汇不匹配。这些现象与之前的工作中得到的结论是一致的。融入外部知识的单一模型(如KAN、CKANN和CKANN-L)优于传统的单一模型和基于BERT的模型。因为它们可以从外部知识和知

识图(KG)中提取相关信息,丰富语义信号,验证了融入外部知识的有效性。同时,本发明可以看到ADDAX的性能明显优于insuranceQA数据集中的几乎所有基线(Test2上的CKANN除外)。

[0133] wikiQA上的结果。从表2中,本发明分析了总共17种方法在wikiQA上的MAP和MRR性能。本发明可以观察到,与一些单一模型(例如,MULT和Multihop-Sequential LSTM)相比,利用外部知识的单一模型无法获得明显的优势。例如,MULT在CKANN上的MAP实现了1.13%的性能增益。这种现象的可能原因是:(1)wikiQA培训数据的缺乏导致相关性学习不足;(2)不相关的外部知识的整合可能会产生语义噪声。第二,关于集合模型,$SUM_{BASE,PTK}$、LRXNET在MAP值和MRR值上均优于SD(BiLSIM+TFM)。显然,集成模型比传统的单一模型和具有外部知识的模型具有更好的匹配性能。这一观察结果表明,集成多个模型对于提高泛化能力是至关重要的。第三,与这些最先进的基于BERT的方法相比,本发明观察到BERT-pooling的性能始终比BERT-attention和HAS差。这一观察结果在三个数据集中都是一致的,这表明交互建模在文本匹配中起着重要作用。相比之下,ADDAX相对于wikiQA数据集中的所有基线获得了更好的性能。

[0134] yahooQA的结果。根据表2中的结果,本发明观察到与insuranceQA数据集类似的性能模式。ADDAX在MAP和MRR方面明显优于所有基线。具体而言,与CKANN(最佳基线)相比,本发明提出的ADDAX的MAP值提高了3.23%。

[0135] 表3.MS MARCO的实验结果。最佳性能以粗体突出显示。▲%表示ADDAX相对于所有基线模型的相对改进。

[0136]

Model	MS MARCO MRR@10(dev)
BM25	18.70
KNRM	19.80
fastText+ConvKNRM	29.00
Duet	24.30
BERT-base	34.70
DeepCT	24.30
docT5query	27.70
ColBERT	34.90
TCT-ColBERT	33.50
COIL-tok	33.60
COIL-full	34.80
CLEAR	33.80
ADORE+ANCE	34.10
ADORE+STAR	34.70
ADR-BERT	36.15
▲%	1.2-7.4

[0137] MS MARCO的实验结果。表3报告了MS MARCO上不同文档检索模型的性能对比结果。从表3中可以发现,首先,传统的查询文档匹配技术(即BM25)的性能始终比深度学习解决方案(如KNRM和fastText+ConvKNRM)差得多,这并不奇怪。其次,对于所有的神经匹配模型,基

于预训练语言模型的方法(例如,BERT-base、ColBERT和COIL-full)比KNRM、fastText+ConvKNRM和Duet获得更好的匹配精度。这是因为预训练的语言模型强大的语言表达能力在很大程度上缓解了词汇不匹配问题。请注意,DeepCT和DocT5Query虽然可以利用预训练语言模型打破术语频率的限制,但它们在语义匹配方面仍然较差。此外,值得注意的是,密集检索器几乎与基于预训练语言模型的模型相媲美。第三,ADDAX始终在MS MARCO数据集上实现最佳性能。ADDAX相比所有基线上的MRR@10增加了1.2%-17.4%。总的来说,在两个不同的任务和数据集上进行的上述比较一致地表明,本发明提出的ADDAX总体上实现了显著的性能提升。这些的结果验证了ADDAX中使用的自适应匹配孪生细胞和哈希去噪器在执行特征识别和去噪可以提高非对称文本匹配精度。

[0138] 4.为验证本发明模型中每一模块对于整体有效,设计了以下对比及消融实验。更具体地说,本发明将ADDAX与以下变体进行了比较:(a)w/o MAG,去除自适应匹配孪生细胞;(b)w/o FD,无公式5-9中所述的特征识别;(c)w/o-HW,省去了高速网络对两种语义信号的融合,而直接相加;(d)不带HD,不包括局部约束哈希去噪器。

[0139] 表4.消融实验结果

[0140]

Model	MS MARCO	wikiQA	
	MRR@10(dev)	MAP	MRR
w/o MAGS	34.90	73.73	74.90
w/o DG	35.76	77.20	78.77
w/o HW	35.32	78.87	80.34
w/o HD	35.49	78.52	80.17
ADDAX	**36.15**	**82.50**	**83.38**

[0141] 表4报告了这些实验在MS MARCO和wikiQA数据集上的实验结果。本发明可以看到,排除自适应匹配孪生细胞会导致最大的性能下降,其次是哈希去噪器。特别是,就MAP值和MRR值而言,wikiQA上的w/o MAG分别下降了8.77%和8.48%,而MS MARCO的MRR@10值下降了1.25%。这表明自适应匹配孪生细胞在ADDAX识别鉴别特征以提高匹配精度中起着至关重要的作用。此外,w/o HD也会导致性能下降,这说明了在文档端执行特征级去噪的有效性。更具体地说,对于MAGS中设计的每个特定结构,本发明还可以得到以下结论:(i)可以观察到w/o FD的性能下降,这表明自适应地凸显不同种类的语义信号是重要的;(ii)w/o HW的性能在一定程度上也有所下降。这表明,高速网络更有效地合成混合鉴别特征。

[0142] 5.超参数的灵敏性分析.此部分,本发明在wikiQA的测试集上对四个重要的超参数(α、δ、γ和h)进行了敏感性分析。从图3中,本发明可以看到,通过将α增加到5(参考图3(c)),可以通过学习更健壮的哈希函数来提高匹配性能。此外,图3(b)通过改变δ值绘制了性能折线图。本发明观察到ADDAX对[1-7,1-5]]范围内不敏感,在δ=1e-6时获得更好的匹配精度。图3(a)通过改变γ值绘制了性能折线图。当γ大于或小于0.003时,性能变得更差。

[0143] 为了选择最合适的低维空间h,本发明进行了实验,在{64,128,256,300,512}之间调整h。结果如图3(d)所示。本发明观察到,当h=300时,ADDAX在wikiQA数据集上始终达到更好的匹配精度。当h变小或变大时,ADDAX会出现一定程度的性能下降。这可能归因于较小的h产生的语义信号不足,而较大的值将不可避免地导致模型过度拟合。

[0144] 上面结合附图对本发明的具体实施方式作了详细说明,但是本发明并不限于上述

实施方式,在本领域普通技术人员所具备的知识范围内,还可以在不脱离本发明宗旨的前提下作出各种变化。

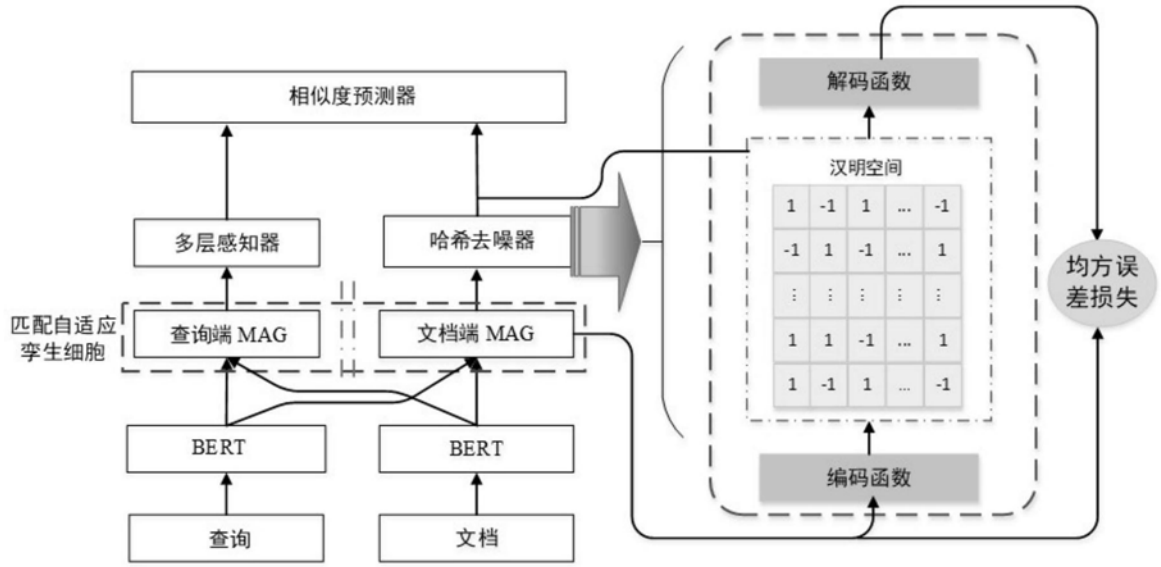

相似度预测器
多层感知器
哈希去噪器
匹配自适应
孪生细胞
查询端 MAG
文档端 MAG
BERT
BERT
查询
文档
解码函数
汉明空间
编码函数
均方误
差损失

图1

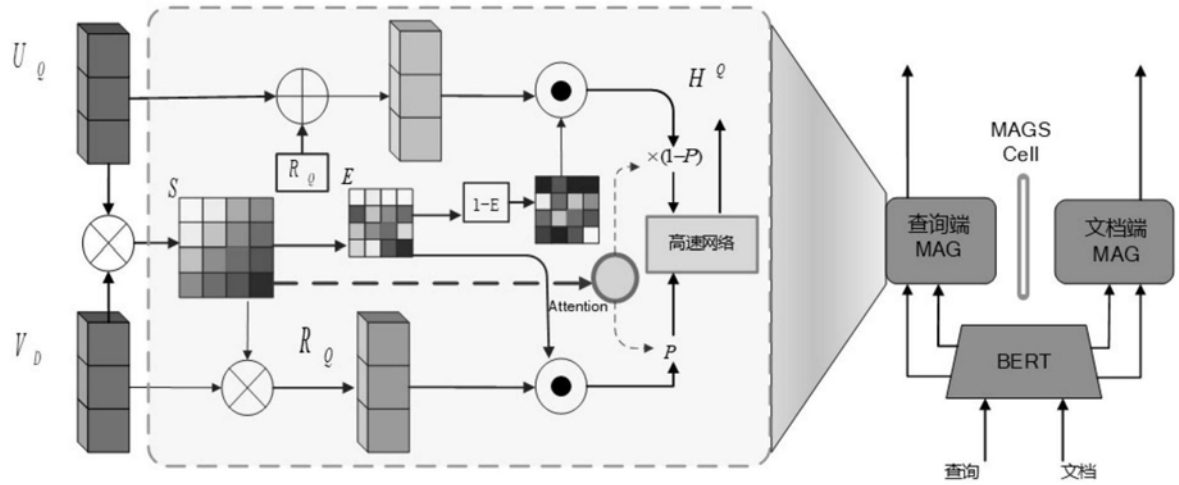
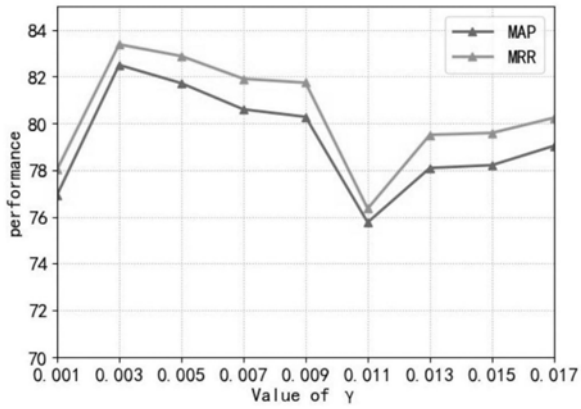
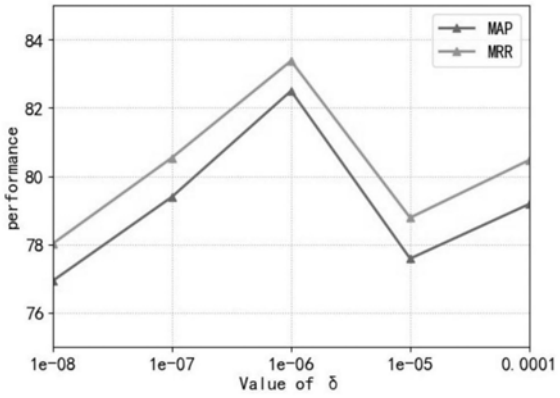

Attention
高速网络
MAGS
Cell
查询端
MAG
文档端
MAG
BERT
查询
文档

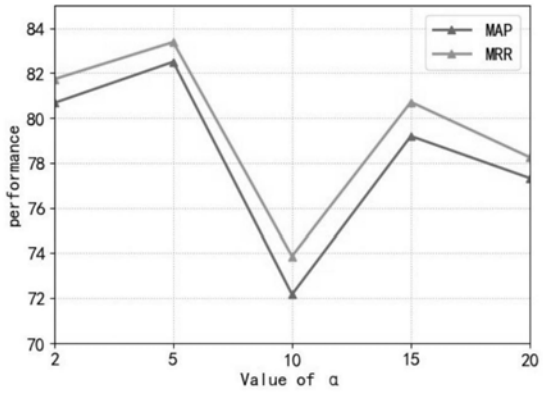
图2


MAP
MRR
performance
Value of γ
0.001 0.003 0.005 0.007 0.009 0.011 0.013 0.015 0.017

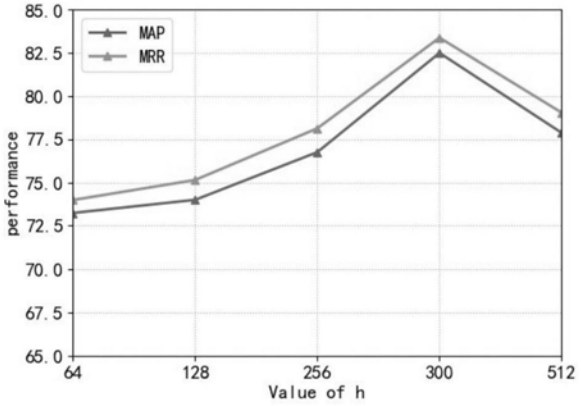
(a)


MAP
MRR
performance
Value of δ
1e-08 1e-07 1e-06 1e-05 0.0001

(b)


MAP
MRR
performance
Value of α

(c)


MAP
MRR
performance
Value of h
64 128 256 300 512

(d)

图3